The effect of vowel height on Voice Onset Time in stop consonants in CV sequences in spontaneous Danish

Mortensen, Johannes; Tøndering, John

Published in:
Proceedings of Fonetik 2013, The XXVIth Annual Phonetics Meeting 12–13 June 2013, Linköping University, Linköping, Sweden

Publication date:
2013

Document version
Early version, also known as pre-print

Citation for published version (APA):
Mortensen, J., & Tøndering, J. (2013). The effect of vowel height on Voice Onset Time in stop consonants in CV sequences in spontaneous Danish. In R. Eklund (Ed.), *Proceedings of Fonetik 2013, The XXVIth Annual Phonetics Meeting 12–13 June 2013, Linköping University, Linköping, Sweden: Studies in Language and Culture no. 21* (pp. 49-52). Linköping University, Department of Culture and Communication. Studies in Language and Culture

The effect of vowel height on Voice Onset Time in stop consonants in CV sequences in spontaneous Danish

Johannes Mortensen & John Tøndering
Department of Scandinavian Studies and Linguistics, University of Copenhagen, Denmark

Abstract

Voice onset time has been reported to vary with the height of vowels following the stop consonant. This paper investigates the effects of vowel height on VOT in Danish CV sequences with stop consonants in Danish spontaneous speech. A significant effect of vowel height on VOT was found in the unaspirated stops [b̥ d̥ g̊], but not in the aspirated stops [p^h t^s k^h]. This is contrary to previous findings.

Introduction

The purpose of this paper is to investigate the effect of vowel height on voice onset time (VOT). Earlier studies have found that VOT before high vowels is significantly longer than VOT before low vowels (Barry & Moyle, 2011; Bijankhan & Nourbakhsh, 2009; Esposito, 2002; Fischer-Jørgensen, 1980; Higgins, Netsell & Schulte, 1998). These previous studies are all based on readings of carefully designed material (read aloud speech), mainly consisting of nonsense words. Krull (1991) compared VOT in Swedish laboratory speech with spontaneous speech. She found no significant differences between VOT in the two speaking styles.

In the present study we have examined CV sequences with stop consonants and stressed vowels in a data set of Danish spontaneous speech. As a measure of vowel height the physiologically defined degree of vowel opening is used. The basic research question to be answered is whether the same relationship between VOT and vowel height exists in spontaneous speech as in read aloud speech.

Note on the Danish stop consonants

The Danish stop consonants can be divided into two series: /ptk/ which are voiceless aspirated stops (except that /t/ is affricated and aspirated) and /bdg/ which are voiceless unaspirated stops. The aspiration distinction is present only in syllable initial position before vowels and sonorant consonants. In medial position the stops are often pronounced as weakly voiced [b d g] (Fischer-Jørgensen & Hutters, 1981). This study, however, is only concerned with syllable initial stops followed by stressed vowels. In this position the stops are pronounced [p^h t^s k^h] and [b̥ d̥ g̊]. In the following the simplified transcriptions [ptk] and [bdg] will be used.

Material

The speech material was taken from the *Danish Phonetically Annotated Spontaneous Speech Corpus* (DanPASS, Grønnum, 2013). The recordings where made at The University of Copenhagen in 1996 and 2004. The speakers were given specific map tasks to talk about in both dialogues and monologues. Although the recordings where made in a studio, the corpus represents an approximation to speech in a natural setting.

The total duration of the recordings is more than 9 hours. The material contains recordings of 27 speakers, 10 women and 17 men. The number of tokens of relevant CV-sequences – i.e. a stop consonant followed by a stressed vowel – exceeds 7,000, but only a little more than 3.000 tokens were included in the present investigation. All tokens in the corpus with [ptk] were included (1.967), whereas tokens with [bdg] were taken from 5 randomly selected speakers (1.074). This procedure was chosen because a very clear pattern was found for tokens with [bdg], whereas this was not the case for [ptk].

Method

Segmentation

Measuring VOT is, to a large extent, dependent on the principles of delimitation. In determining the starting and ending point of VOT, the recommendations in Fischer-Jørgensen and Hutters (1981) were followed. Accordingly, VOT starts at the first release of the closure, which can be located in the waveform, and ends with the start of the higher formants of the following vowel, which can be located in the spectrogram. There is an uncertainty of some milliseconds in both points.

Measurements were only made in stressed syllables, and in case of doubt about the segmentation – e.g. if the release of the closure was so smooth that it could not be located in the waveform – no measurement was made.

Vowel classification

Apart from using spontaneous speech, another new methodological approach in this study is the use of physiological vowel description. In the traditional vowel classification the distance from tongue to palate defines vowel height, but in the physiological description vowel height is defined as the degree of constriction at the narrowest point in the vocal tract (Grønnum, 2005). We have chosen to use the physiological vowel description because the most prevailing explanation for the effect of vowel height on VOT is related to the flow of air through the vocal tract. This airflow does not depend (directly) on tongue height, but more precisely on the degree of constriction, or conversely on the degree of opening: One of the conditions for the vocal folds to start vibrating – i.e. voice onset – is that the air pressure above the glottis is lower than the air pressure below the glottis. During the articulation of a stop consonant, however, the supraglottal pressure will rise because the flow of air is restrained at the point of constriction so that the air builds up behind this point – the place of articulation of the stop consonant. After the release of the closure the air will flow rapidly out of the vocal tract and the pressure just above glottis will start to fall again. It will take a little time for the air pressure to drop sufficiently for the vocal cords to start vibrating. How long it takes will depend on the size of the passage through the mouth. If the passage is wide the pressure will drop quickly, but if the passage is narrow the pressure will drop more slowly. The consequence of this is that the narrow passage in high vowels delays the airflow through the mouth and thereby also delays voice onset and increases VOT (Fischer-Jørgensen, 1980). The physiological classification of the Danish vowels follows Grønnum (2005:105), except from merging her levels 4 and 5, resulting in four levels on the scale of degree of opening. The classification of the stressed Danish vowels can be seen from *Table 1*. Level 1 is the narrow end of the scale and level 4 is the open end (notice that the terms high and low makes less sense here, although high would correspond to level 1 and low would correspond to level 4). The merging of level 4 and 5 is done partly because the vowels in these levels are rather infrequent resulting in fewer tokens in the material, partly because some of them can be suspiciously difficult to distinguish.

Table 1. Physiological classification of the degree of opening of Danish vowels. The phonetic transcription is according to the modified version of the IPA for Danish (Grønnum, 2005).

Degree of opening			
1	2	3	4
[i] [y] [u] [o] [ɒ]	[e] [ɔ] [ʌ]	[ɛ] [ø] [ɑ]	[æ] [œ] [œ̞] [a] [ɶ]

A drawback of using this physiological classification of the Danish vowels is that it is based on a combination of x-rays of vowel articulation of one person and introspection (Grønnum, 2005). A proper physiological investigation on this aspect of the Danish vowels does not exist. However, since this study has no intention of detecting minor fluctuations in VOT, but only intends to reach conclusions regarding the overall tendency in the relation between vowel height and VOT, Grønnum's more or less introspective classification should be an acceptable point of departure.

Statistical approaches

The data was analysed using plots and statistical modeling. Mixed effects multiple regression models were used with speaker and word as random effects and the degree of vowel opening as fixed effect. Models with additional variables – such as gender of the speaker – were also tried out, but this did not yield particularly interesting results and will not be reported here. Each stop consonant was analysed separately and after fitting the models, residual diagnostics were carried out to validate model assumptions.

It should be noted, that in order to be able to use regression modeling, the degree of vowel opening was treated as a continuous variable although it strictly speaking is only ordinal. It is reasonable to think that Grønnum (2005) has chosen the levels of vowel height so that the distance between them is fairly constant, but as noted earlier the classification is partly introspective and also the two highest levels are taken as one in this study.

Results

Visual inspection

Figure 1 and *Table 2* show the mean VOT for the different stops according to vowel height. In [bdg] there is a clear tendency for a fall in VOT when the degree of vowel opening increases. The only exception is a rise in [b] from level 3 to 4. The difference in the average VOT of the narrowest vowels and the most open vowels is a bit less than 10 ms.

Table 2. Mean VOT in ms.

Stop consonant		Degree of vowel opening			
	No.	1	2	3	4
[b]	156	19.2	16.3	11.7	14.0
[d]	608	29.2	24.6	20.5	18.0
[g]	310	34.0	29.4	28.2	25.1
[p]	341	77.0	64.0	71.9	58.1
[t]	600	86.6	87.2	81.6	88.7
[k]	1.026	81.0	61.6	60.5	76.1

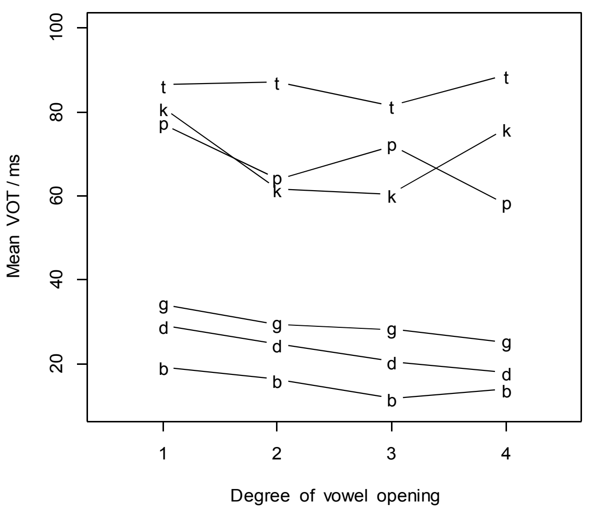

Figure 1. Mean VOT as a function of vowel opening degree.

In [ptk] the picture is less clear. In [k] there is a steep fall from level 1 to 2, then almost no difference from 2 to 3, and then a steep rise to level 4. In [p] the pattern is falling–rising–falling. In [t] the variation seems to be rather small although there is a slight fall from level 2 to 3 and a somewhat larger rise from level 3 to 4. Overall there does not seem to be any consistent effect of vowel height on the VOT of [ptk].

It should be noted that the variation in VOT between individual tokens of CV sequences is larger than the variation across different vowel heights, so that for instance a [d] in level 4 can easily have a longer VOT than a [d] in level 1. This is not apparent in *Figure 1* because it only shows the mean values. Still it is clear that at least in [bdg] vowel height accounts for some of the variation in VOT.

Statistical analysis

Table 3 shows coefficients, standard deviations and p-values for the mixed effects models with degree of vowel opening as the fixed effect. In [b], [d] and [g] the effect of vowel height is significant with p-values below 0.001. In [p], [t] and [k] the effect of vowel height is not significant.

Table 3. Result of the statistical analysis with degree of vowel opening as fixed effect.

Stop consonant	Coefficient	Standard deviation	*p*-value
[b]	-0.173	0.040	< 0.001
[d]	-0.166	0.019	< 0.001
[g]	-0.104	0.018	< 0.001
[p]	-1.766	2.345	0.452
[t]	-0.014	0.016	0.376
[k]	-0.031	0.020	0.114

All the coefficients are negative indicating a fall in VOT when the degree of opening increases. This is also the case in [ptk] although the coefficients here are very small and the effect is insignificant. The reason for the seemingly high values in [p] is technical: the VOT values of the other stop consonants were log-transformed because their distributions were skewed, which was not the case for [p].

Supplementary analysis

To check whether the lack of effect of vowel height on VOT in aspirated stops could be due to the use of the physiologically based classification of vowel height, we also examined the association between VOT and vowel height according to the traditional description. In this analysis the Danish vowels were classified in three degrees as high ([i y u e]), mid ([ɛ ø o œ ɔ]) or low ([a ɶ ʌ ɑ æ œ̞ ɒ]). The mean VOT divided by vowel height is shown in *Table 4*. As seen, the overall pattern is the same as when using physiological vowel classification.

Table 4. Mean VOT in ms divided by vowel height.

Stop consonant	Vowel height		
	High	Mid	Low
[b]	19.0	18.0	12.7
[d]	25.9	21.0	18.4
[g]	36.0	28.8	27.4
[p]	70.1	65.0	67.8
[t]	86.6	86.4	83.8
[k]	80.1	62.7	66.3

Discussion and conclusion

The effect of vowel height on VOT was investigated and for [bdg] a clear relation was found. VOT was shorter when the vowels were more open. This result is in agreement with previous findings (Fischer-Jørgensen, 1980, and others). For [ptk], however, no significant effect of vowel height on VOT was found. This is contrary to previous findings and calls for discussion.

The most obvious explanation might be methodological. The use of physiological vowel description could be the reason for the deviating result. However, the supplementary analysis based on the traditional vowel description yielded nearly the same results – ruling out the explanation of difference in vowel description.

The use of spontaneous speech might be another explanation, but it is hard to explain why this should have an effect in [ptk], but not in [bdg].

The deviating results could also be a consequence of differences in the criteria used for the delimitation of VOT. As described in Fischer-Jørgensen and Hutters (1981) it is possible to consider the vowel to start at different points, and the choice of point might have serious consequences for the length of VOT, especially before low vowels. Fischer-Jørgensen and Hutters (1981) argues in favour of choosing the start of the higher formants as the vowel onset. This suggestion is followed in the present study, but most studies take the start of voicing as the endpoint of VOT, which does not seem unreasonable when dealing with the concept voice onset time (Barry & Moyle, 2011; Bijankhan & Nourbakhsh, 2009).

There might, however, also be another explanation for the lack of effect of vowel height on VOT in [ptk] found in this study. As noted earlier, the most prevailing explanation in the literature for the effect of vowel height on VOT is related to the flow of air through the vocal tract. But the question is, whether this explanation holds when VOT is relatively long. In other words, maybe the VOT in [ptk] is so extensive that the sufficient oral pressure drop is reached long before voice onset – regardless of the height of the following vowel. If this is the case, vowel height might not have much effect on VOT in [ptk]. One way to investigate this could be to examine if the effect of vowel height on VOT in [ptk] is present for some speakers and not for others, and whether this possible pattern would be related to speakers's overall level of voice onset time. This will be the subject for further studies.

References

Barry, J. & M. Moyle. 2011. Covariation among vowel height effects on acoustic measures. *Journal of the Acoustical Society of America*, 130(5): 365−371.

Bijankhan, M. & M. Nourbakhsh. 2009. Voice onset time in Persian initial and intervocalic stop production. *Journal of the International Phonetic Association* 39:335−364.

Esposito, A. 2002. On vowel height and consonantal voicing effects: Data from Italian. *Phonetica* 59:197−231.

Fischer-Jørgensen, E. 1980. Temporal relations in Danish tautosyllabic CV sequences with stop consonants. *ARIPUC* 14:207−261.

Fischer-Jørgensen, E. & B. Hütters. 1981. Aspirated stop consonants before low vowels. A problem of delimitation. *ARIPUC* 15:77–102.

Grønnum, N. 2013. Danish Phonetically Annotated Spontaneous Speech.

Grønnum, N. 2005. *Fonetik og Fonologi. Almen og Dansk.* Copenhagen: Akademisk Forlag.

Higgins, M. B., R. Netsell & L. Schulte. 1998. Vowel-related differences in laryngeal articulatory and phonatory function. *Journal of Speech, Language and Hearing Research* 41(4):712−724.

Krull, D. 1991. VOT in spontaneous speech and in citation form words. Phonetic Experimental Research, Institute of Linguistics, University of Stockholm (*PERILUS*). No. XII:101−107.

Proceedings of Fonetik 2013

The XXVIth Annual Phonetics Meeting
12–13 June 2013, Linköping University
Linköping, Sweden

Studies in Language and Culture
no. 21

Robert Eklund, editor

Linköping University

Conference website: www.liu.se/ikk/fonetik2013
Proceedings also available at: http://roberteklund.info/conferences/fonetik2013

Cover design and photographs by Robert Eklund
Photo of Claes-Christian Elert taken by Eva Strangert on the occasion of his 80th birthday

Proceedings of Fonetik 2013, the XXVIth Swedish Phonetics Conference
held at Linköping University, 12–13 June 2013
Studies in Language and Culture, no. 21
Editor: Robert Eklund
Department of Culture and Communication
Linköping University
SE-581 83 Linköping, Sweden

ISBN 978-91-7519-582-7
eISBN 978-91-7519-579-7
ISSN 1403-2570

Printed by LiU-Tryck, Linköping, Sweden, 2013